Spectral Algorithms & SoS

Announcements

1. HW3 goes out today.
2. SoS Hats!

Average-case problems

2 broad categories.

1. Refutation Problems.

   input instances ~ D.

   whp the instance is a NO instance "null"

   Goal: efficiently find a certificate of there being no sol^n.

ex. refuting random CSPs

2. "Planted" problems.

   instances ~ D.

   whp, instance is a YES instance

   Goal: find a solution.

   e.g. tensor decomposition ✓

   Gaussian Mixture clustering ✓

3. Distinguishing problem

$D_{planted}$ (YES) $\qquad$ D_{null} (NO)

input instance is either $\sim D_{planted}$
or $\sim D_{null}$

Goal: decide correctly whp which dist^n generated the input!

Example 1: max-clique problem.

Clique: $S \subseteq [n]$ such that all pairs are conn in G

D_{null}: $G \sim G(n, \frac{1}{2})$.

Q1: What's the size of max-clique in G?

Ans: $2 \cdot \log_2(n)$. (Bollobás)

"two-point-conc^n"

Sketch for why: $|S| = t$

Chance that S is a clique: $2^{-\binom{t}{2}}$.

expected # of cliques of size t in G $= \binom{n}{t} \cdot 2^{-\binom{t}{2}}$.

$\sim n^t \cdot 2^{-\frac{t^2}{2}}$

$t!$

$\sim 2^{+t \cdot \log n - t \cdot \log t - \frac{t^2}{2}}$

< 1 if $\frac{t^2}{2} \sim t \log(n)$

$t \sim 2 \cdot \log_2 n$

<u>Conclusion:</u> max-clique in G

$\leq O(\log_2(n))$ w.p. 0.99

<u>Solution:</u> clique of size $> 3 \cdot \log_2(n)$.

then $G \sim G(n, \frac{1}{2})$ doesn't have such a

sol^n.

<u>Refutation Problem</u>: Give an algo that

takes input G on n vertices.

1. outputs $V \in \mathbb{N}$.
2. <u>Correctness</u>: $V \geq$ max-clique(G).
3. $\Pr_{G \sim G(n, \frac{1}{2})} [V \leq \omega] \geq 0.99$

for as small an ω as possible

<u>Question (Refutation)</u>: For what ω does there exist an efficient (poly(n). time) ref algo for max-clique in G.

<u>Note</u>: $\omega \geq 2 \cdot \log_2(n)$

<u>brute-force algo</u>: 1) go over all subsets of size ω

2) check if there's a clique in them.

runs in time $n^{O(\log(n))}$.

<u>Next</u>: we'll study a "$\underbrace{\text{Spectral Refutation}}$" algorithm

<u>Spectral Refutation</u>: $G \sim G(n, \frac{1}{2})$.

① build a matrix $M(G)$ in poly time
$n^d \times n^d$

(2) relate the value of max-clique to $\boxed{\text{largest}}$ eigenvalue of $M(G)$.

(3) output $\|M(G)\|_2$

In general. Spectral algo builds a matrix out of input efficiently & computes the largest eigenvalue

Spectral Refutation for Max-Clique in $G(n,\frac{1}{2})$:

1) What matrix?

$A \in \mathbb{R}^{n \times n}$, $A(i,j) = \begin{cases} +1 & \text{if } i \sim j \text{ in } G \\ -1 & \text{if } i \not\sim j \text{ in } G \end{cases}$

2) Suppose $x \in \{0,1\}^n$ [indicator of a subset of vertices]

Obsn 1:
Suppose x indicates a clique in G.

$$x^T A x = \sum_{ij} x_i x_j A_{ij}$$

$$= \left(\sum_i x_i\right)^2 - \left(\sum_i x_i^2\right) \quad \text{— } (*)$$

obs^n 2:

$$x^T A x \leq \|x\|_2^2 \cdot \|A\|_2 \quad \text{— } (**)$$

$(*) + (**)$

$$\hookrightarrow \left(\sum_i x_i^2\right)^2 - \left(\sum_i x_i^2\right) \leq \|x\|_2^2 \cdot \|A\|_2$$

$$\sum_i x_i^2 - 1 \leq \|A\|_2$$

or $$\|x\|_2^2 \leq \|A\|_2 + 1$$

(size of clique)

<u>Conclusion:</u> $\|A\|_2 + 1$ is an upper bound on max-clique in G.

<u>Ref algo:</u> 1) return $\|A\|_2 + 1$

<u>Analysis:</u> how good (i.e. how small is the typical ω) is the algo?

Claim: Suppose $G \sim G(n, \frac{1}{2})$. Then w.p. 0.99 $\|A\|_2 \lesssim O(\sqrt{n})$.

Proof: HWO. (random matrix thy)

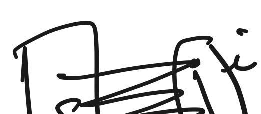
Fact: Suppose $M \in \mathbb{R}^{n \times n}$, symmetric, all entries in upper triangular part of M are indep random. Suppose further that each such entry has mean 0, variance ≤ 1 & $\mathbb{E}\, M_{i,j}^4 \lesssim O(1)$. Then $\|M\|_2 = O(\sqrt{n})$ w.p. $1-o(1)$. ✓

Please read Tao's blog-post.

Conclusion: There's a ref algo for max-clique on $G(n, \frac{1}{2})$ with $\omega = O(\sqrt{n})$.

$\boxed{\text{"Information Computation Gap"}}$

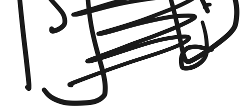
A ... algorithm :



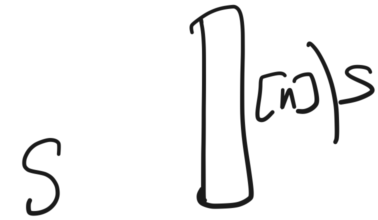
An improved algo

$S \subseteq [n]$, $|S| = t$, $A_S \in \mathbb{R}^{n-t \times n-t}$

$$A_S(i,j) : \begin{cases} +1 & \text{if } i \sim j \ \& \ \{i,j\} \in \text{comm-nbr}(S) \\ -1 & i \nsim j \ \& \ \{i,j\} \in \text{comm-nbr}(S). \\ 0 & \text{o/w.} \end{cases}$$



Suppose x is a clique in G contains vertices in $S \subseteq [n]$ of size $|S| = t$.



Then: $\underline{x^T \cdot A_S \cdot x} =$

$$\sum_{i,j \in [n]\setminus S} x_i \cdot x_j \cdot A_S(i,j)$$

= # edges betⁿ all pairs of vertices in set indicated by x minus S.

$$\sim \frac{(\|x\|_2^2 - |S|)^2}{2}$$

$$\frac{(\|x\|_2^2 - |S|)^2}{2} \le x^T A_S x \le \|x\|_2^2 \cdot \|A_S\|_2$$

$$\frac{1}{2}\|x\|_2^4 + \frac{1}{2}|S|^2 - \|x\|_2^2 \cdot |S| \lesssim \|x\|_2^2 \cdot \|A_S\|_2$$

$$\frac{1}{2}\cdot\|x\|_2^2 + \frac{1}{2}\cdot\frac{|S|}{\|x\|_2^2} - |S| \leq \|A_S\|_2$$

$$\|x\|_2^2 \leq 2\cdot\|A_S\|_2 + 2|S|.$$

<u>thus</u> $\forall$ clique indicators:

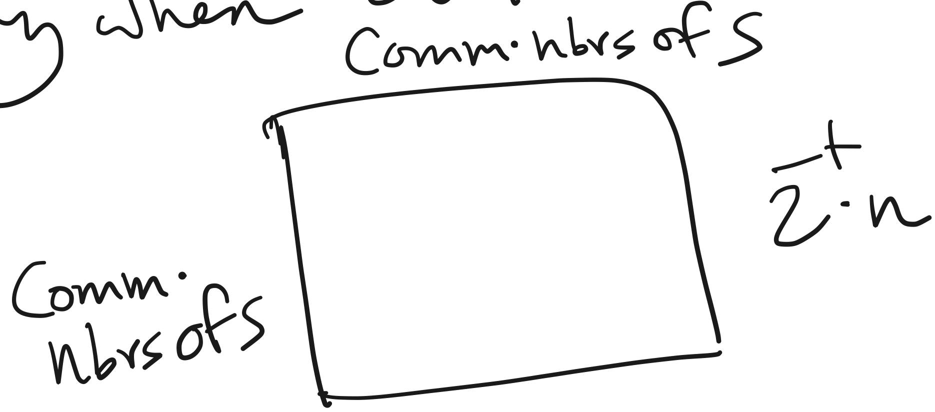
$$\|x\|_2^2 \leq \underbrace{2\cdot\max_{|S|=t}\|A_S\|_2 + 2\cdot|S|}.$$

We've constructed $\binom{n}{t}$ diff matrices

<u>Question:</u> is this a better upper bound?

When $A_S(i,j)$ non-zero?

only when $S\cup\{i\}$ is a clique: 2^{-t} chance



Applying Fact:

$$\|x\|_2^2 \leq \underset{\sim}{2t} + \underbrace{O\left(\sqrt{\frac{n}{2^t}}\right)}.$$

For $t \ll \log_2(n)$:

($t = O(1)$ for poly time algos)

<u>Conclusion:</u> $\forall$ constants $c > 0$ $\exists$ a poly time refutation algo (with $\omega \leq c \cdot \sqrt{n}$).

* <u>Matrix Chernoff bound</u>:

Let $B_1, B_2, \_\_\_ \to \mathbb{R}^{n \times n}$ fixed matrices such that $\forall i : \underline{\|B_i\|_2} \leq O(1)$ & $\underline{\|\sum_i B_i^2\|_2} \leq \sigma^2$

Let $r_1, r_2, \_\_\_$ be independent $\{\pm 1\}$

<u>Then</u>, $\Pr\left[\|\sum_i r_i B_i\|_2 \geq t\right] \leq 2 \cdot \underline{n} \cdot e^{-\frac{t^2}{2\sigma^2}}$.

Sketch: Proof of weaker bound in Fact

Using matrix Chernoff.

$$A = \sum_{i,j} \underbrace{A_{i,j}} \cdot e_i e_j^T$$

apply M.C. bound with $\gamma_i = A_{i,j}$s

& B_is equal to $e_i e_j^T$

① $\|e_i e_j^T\|_2 \leq 1$. (HW0)

② $\|\sum_{i,j} (e_i e_j^T)^2\|_2 = n$.

Can obtain a bound of $O(\sqrt{n} \cdot \log(n))$.

Spectral algorithm is "captured" by SOS.

* SOS for refuting max-clique.

x sat $\mathcal{A}$ iff x $\left\{ \begin{array}{ll} x_i^2 = x_i & \forall i \\ x_i x_j = 0 & \forall i,j : i \nsim j \end{array} \right\} \mathcal{A}$

is a clique in G

SoS-ref-algo: output $\max_{\mu \text{ is a p.d. sat } \mathcal{A}} \tilde{\mathbb{E}}_\mu \sum_i x_i$

① Why is this ref algo? ✓

② What's the typical value of this program when $G \sim G(n, \frac{1}{2})$.

For degree 2: typical val $\sim O(\sqrt{n})$.

Lemma:

$$\mathcal{A} \vdash_X^4 \left\{ \|x\|_2^4 \leq \left(\sum_i x_i^2\right) + \|x\|_2^2 \cdot \|A\|_2 \right\}$$

Proof: $\mathcal{A} \vdash$

$$\left(\sum_i x_i\right)^2 = \sum_{i,j} x_i \cdot x_j$$

$$= \left[\sum_i x_i^2 + \underline{\sum_{i \sim j} x_i x_j} - \underline{\sum_{i \not\sim j} x_i x_j}\right]$$

$$+ \left[2 \cdot \underbrace{\sum_{i \not\sim j} x_i \cdot x_j}_{i \neq j}\right]$$

$$= \sum_i x_i^2 + x^T A x$$

$$\le \sum_i x_i^2 + \underline{\|x\|_2^2 \|A\|_2}$$

$\underline{\forall M}$. $\quad \vdash_x^2 \{ x^T M x \le \|M\|_2 \cdot \|x\|_2^2 \}$

$\underline{\text{Pf}}$: $\Leftrightarrow x^T \underbrace{(\|M\|_2 \cdot \mathbb{I} - M)}_{\text{4 ps}} x \succeq 0$

$$V^T V$$

$$\|$$

$$\|Vx\|_2^2$$

SoS proof. $\boxed{\|M\|_2 \cdot \|x\|_2^2 - x^T M x = \|Vx\|_2^2}$

$\underline{\text{Now}}$: $\tilde{\mathbb{E}} \|x\|_2^4 \le \tilde{\mathbb{E}}(\sum_i x_i^2) \quad - (*)$

$$+ (\tilde{\mathbb{E}} \|x\|_2^2) \|A\|_2$$

$$\Rightarrow (\tilde{\mathbb{E}} \|x\|_2^2)^2 \le \underline{(\tilde{\mathbb{E}} \cdot \|x\|_2^2)}(1 + \|A\|_2)$$

$\hookrightarrow$ ***

$$\left(\tilde{\mathbb{E}}\|x\|_2^2\right)^2 \leq \tilde{\mathbb{E}}\|x\|_2^4$$ by Cauchy Schwarz

$$\boxed{\text{How? } f=\|x\|_2^2;\ g=1 \qquad \tilde{\mathbb{E}} f\cdot g \leq \sqrt{\tilde{\mathbb{E}} f^2}\sqrt{\tilde{\mathbb{E}} g^2}}$$

$$*** \Rightarrow \boxed{\tilde{\mathbb{E}}\|x\|_2^2 \leq 1+\|A\|_2}.$$

① Can also solve the "planted" problem — by looking at the largest eigenvector of $A(G)$.

Planted Clique: $G \sim G(n, \frac{1}{2})$
$S \subseteq [n]$, $|S| = \omega \gg 2\log_2(n)$
Complete the clique on S in G to obtain G'
input $= G'$.

Alon-Krivilevich + Sudakov '98

② distinguishing problem. (for $\omega \gg \sqrt{n}$)

Dplanted Dnull

[illegible] planted

algo: $\|A(G)\|_2 > \omega$ output planted
& $\|A(G)\|_2 \leq \omega$ output null.

refutation $\Rightarrow$ distinguishing

Example 2: Refuting R-CSPs

4-XOR. $\mathcal{I}$: $\underbrace{x_i x_j x_k \cdot x_e}_{\pm 1\text{-valued vars}} = b_{ijke} \in \{\pm 1\}$

$\mathcal{I}$: pick m 4-tuples + RHS at random & independent
& output the resulting

If $m > O\left(\frac{n}{\varepsilon^2}\right)$ then no more than $\frac{1}{2}+\varepsilon$ frac constraints can be sat by any assignment x.

Thus $m \gg O\left(\frac{n}{\varepsilon^2}\right)$, we've a natural "null" distn

Refutation: input $= \mathcal{I}$
output: v

$v \geq \mathrm{Val}(\mathcal{I})$.

Correctness: ...

Utility: $\Pr_{I \sim D_{null}} [v \leq 1-\delta]$ w.p. 0.99 for some $\delta > 0$

Grigoriev's Thm: $\exists$ $\Omega(n)$-deg pseudo-distn that satisfies all constraints of a random 4-XOR formula with $\boxed{m \leq O(n)}$ constraints w.p 0.99 over the draw of the instance.

Today: we'll give a poly time ref algo that succeeds in proving a random 4-XOR instance unsat w.p. ≥ 0.99 whenever $m \gg O(n^2)$

Let's construct a spectral refutation.

What matrix?

$$A \in \mathbb{R}^{\binom{n}{2} \times \binom{n}{2}}.$$

$$A((i,j),(k,l)) = \begin{cases} \frac{b_{ijkl}}{6} & \text{if } i,j,k,l \text{ appears in constraints} \\ 0 & \text{o/w} \end{cases}$$

" $I(x)$

Consider: $\frac{1}{m}\sum_{i,j,k,\ell}^{\mathcal{I}} x_i x_j x_k x_\ell \cdot A(\{i,j\},\{k,\ell\})$

$= 1$ if x satisfies all constraints

more generally $\mathrm{val}_{\mathcal{I}}(x) \leq \frac{1}{2} + \frac{\mathcal{I}(x)}{2}$

frac of constraints in $\mathcal{I}$ sat by x

to **refute**, must prove: $\mathcal{I}(x) \leq \varepsilon$ $\forall x \in \{\pm 1\}^n$

$$\mathcal{I}(x) = \frac{1}{m} \sum_{i,j,k,\ell} x_i x_j x_k x_\ell \cdot A(\{i,j\},\{k,\ell\})$$

$$= \frac{1}{m} (x^{\otimes 2})^T \cdot A \cdot x^{\otimes 2}$$

$$\leq \frac{1}{m} \cdot \|x^{\otimes 2}\|_2^2 \cdot \|A\|_2$$

$$= \frac{1}{m} \cdot n^2 \cdot \|A\|_2.$$

Ref algo: output $\frac{1}{2} + \frac{1}{2} \cdot \frac{n^2}{m} \cdot \|A\|_2$

Analysis: how good is this upper bound?

For every $C = \{i,j,k,\ell\}$

$$A_C((i,j),(k,\ell)) = \begin{cases} \frac{+1}{24} & \text{if } C=\{i,j,k,\ell\} \\ 0 & \text{o/w.} \end{cases}$$

Then. $A = \sum_C b_C \cdot A_C$

Next: ① $\|A_C\|_2 \le O(1)$. $(\forall C)$.

② $\left\|\sum_C A_C^2\right\|_2 \overset{?}{\lesssim}$

$$A_C^2(r,s,t,u) = \sum_{v,w} A_C((r,s),(v,w)) \cdot A((v,w),(t,u))$$

non-zero only if $\{r,s\} = \{t,u\}$
& $\{r,s\} \subseteq \{i,j,k,\ell\} = C$.

$\implies$ A_C^2 is diagonal

$\boxed{\sum_C A_C^2}$: diagonal matrix.

Fix $\{i,j\}$: Then # constraints

$P(X \text{ any } \{i,j\})$ ...

Containing $\{i,j\}\} = O\left(\frac{m \cdot \log n}{n^2}\right) + O(\log n)$

whp.

Since $\sum A_c^2$ is diagonal.

$$\|\sum A_c^2\|_2 \leq \text{max-entry of } \sum_c A_c^2$$

$$= \underline{O\left(\frac{m \cdot \log(n)}{n^2}\right)} + O(\log n) \rightarrow \sigma^2$$

<u>Matrix-Chernoff:</u>

$$\underline{\Pr[\|\sum_c A_c\| \geq t] \sim \underbrace{2n^2 e^{-\frac{t^2}{\sigma^2}}}_{0.01}}$$

$$t \sim O(\sigma \sqrt{\log(n)})$$

$$\Rightarrow \underline{\|\overset{A}{\overset{\shortparallel}{\sum_c A_c}}\|_2} \leq \tilde{O}\left(\frac{\sqrt{m}}{n}\right) \text{ w.p. } 0.99$$

When $\underline{m \gg n^2}$, 1st term dominates.

ref upper bound:

$$\ldots \leq \ldots + \ldots n^2 \|A\|$$

$$\mathrm{Val}_I \lesssim \frac{1}{2} + \frac{1}{2}\cdot\frac{\cdots}{m}\cdot\|\cdots\|_2$$

$$= \frac{1}{2} + \frac{1}{2}\cdot\frac{n^2}{m}\cdot\frac{\sqrt{m}}{n}\cdot\mathrm{polylog}(n)$$

$$= \frac{1}{2} + \frac{1}{2}\cdot\frac{n}{\sqrt{m}}\cdot\mathrm{polylog}(n)$$

$$\ll \frac{1}{2} + \varepsilon$$

if $m \ggg \hat{O}(n^2/\varepsilon^2)$.

factor n higher than info theoretic min m

"information-computation gap"

no known poly time ref algo that succeeds when $m \leq n^{2-o(1)}$

* Planted Model (m)

1) pick m 4-tuples at random
2) pick $x^* \in \{\pm 1\}^n$ "planted assignment"

3) for every 4-tuple i, j, k, l picked ...

Choose RHS to be $x_i^* x_j^* x_k^* x_l^*$ w.p. 0.99

& $-x_i^* x_j^* x_k^* x_l^*$ w.p. 0.01

Ask: ① Can you find the planted assignment? when $m \ll n^2$? don't know.

If $m \gg n^2$, there's a poly time Spectral algo!

② distinguishing variant:

$m \gg n^2$, spectral algo succeeds.

$\boxed{\text{Can reduce other CSPs to XOR}}$ ✓

* Can "capture" this spectral algo within SOS in a similar way as for clique

---

refutation problem → Spectral algo (essentially best known)

↓

Captured within SOS

planted problems
distinguishing problem $\leftarrow$ spectral/ref algo

Summarize: spectral refutation

$\downarrow$

SoS refutation

$\downarrow$

distinguishing.

Theorem: For a "large-class average-case problems" if there's a degree d SoS refutation then there's a spectral refutation with a $n^{O(d)} \times n^{O(d)}$ matrix whose entries are $O(d)$ deg polynomials in input.

$\rightarrow$ in general, this thm is non-constructive

$\rightarrow$ but can come up with the right matrix explicitly in natural problems.

Tensor decomposition:

SoS based algo $\xrightarrow[\text{existential}]{\text{generic}}$ spectral algo with same performance guarantees

Speed-up under additional assumptions.